UNIVERSITY OF ILLINOIS BULLETIN

ISSUED WEEKLY

VOL. XXV SEPTEMBER 13, 1927 No. 2

[Entered as second-class matter December 11, 1912, at the post office at Urbana, Illinois, under the Act of August 24, 1912. Acceptance for mailing at the special rate of postage provided for in section 1103, Act of October 3, 1917, authorized July 31, 1918.]

BULLETIN NO. 37

BUREAU OF EDUCATIONAL RESEARCH
COLLEGE OF EDUCATION

PREDICTING THE SCHOLASTIC SUCCESS OF COLLEGE FRESHMEN

By

CHARLES W. ODELL
Assistant Director, Bureau of Educational Research

PRICE 25 CENTS

PUBLISHED BY THE UNIVERSITY OF ILLINOIS, URBANA
1927

The Bureau of Educational Research was established by act of the Board of Trustees June 1, 1918. It is the purpose of the Bureau to conduct original investigations in the field of education, to summarize and bring to the attention of school people the results of research elsewhere, and to be of service to the schools of the state in other ways.

The results of original investigations carried on by the Bureau of Educational Research are published in the form of bulletins. A list of available publications is given on the back cover of this bulletin. At the present time five or six original investigations are reported each year. The accounts of research conducted elsewhere and other communications to the school men of the state are published in the form of educational research circulars. From ten to fifteen of these are issued each year.

The Bureau is a department of the College of Education. Its immediate direction is vested in a Director, who is also an instructor in the College of Education. Under his supervision research is carried on by other members of the Bureau staff and also by graduates who are working on theses. From this point of view the Bureau of Educational Research is a research laboratory for the College of Education.

BUREAU OF EDUCATIONAL RESEARCH
College of Education
University of Illinois, Urbana

BULLETIN NO. 37

BUREAU OF EDUCATIONAL RESEARCH
COLLEGE OF EDUCATION

PREDICTING THE SCHOLASTIC SUCCESS OF COLLEGE FRESHMEN

By

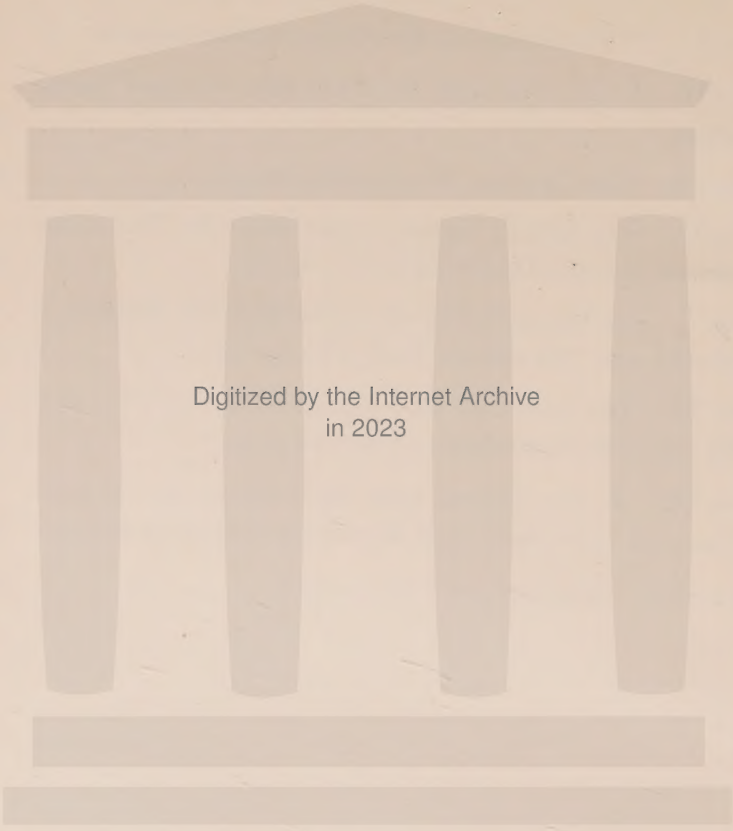
CHARLES W. ODELL
Assistant Director, Bureau of Educational Research

PRICE 25 CENTS

PUBLISHED BY THE UNIVERSITY OF ILLINOIS, URBANA
1927


UNIVERSITY
OF ILLINOIS
PRESS

TABLE OF CONTENTS



PREFACE

The prediction of the scholastic success of college freshmen is commanding the attention of many persons, especially those who are responsible for the administration of colleges and universities. It has been proposed that by making use of a student's high-school record and by administering an intelligence test it would be possible for an institution to predict the probable success of an entering student in the various subjects of instruction. It is, of course, generally recognized that such predictions would not be accurate in all cases. Some authorities contend that, if sufficient information is secured and the prediction is made intelligently, it will be sufficiently accurate to be very helpful in guiding the student when he enters college. Other authorities maintain that in general the prediction will be so inaccurate that it will not be very useful.

In this bulletin Dr. Odell presents the results of a very careful inquiry into the accuracy of predictions that may be made using certain information. Although knowing the probable accuracy of the predictions that may be made does not determine the value of such predictions, it should be helpful to college administrators to know the probable accuracy of the predictions they may make. Hence, it is believed that Dr. Odell's study constitutes a significant contribution in the field of college administration.

Walter S. Monroe, *Director*

May 4, 1927

PREDICTING THE SCHOLASTIC SUCCESS OF COLLEGE FRESHMEN

CHAPTER I

INTRODUCTION AND STATEMENT OF THE PROBLEM

The recent increase in college enrollment and two resulting problems. In a previous publication[1] the present writer has called attention to the fact that one of the most notable and significant recent educational tendencies in this country has been the marked increase in school enrollment, especially on the secondary and higher levels. Such a tendency has, it is true, existed from the beginning of our educational system, but, since sometime near the end of the Nineteenth Century and even more since the close of the World War, it has been greatly accentuated. This may be seen from the last Statistical Summary of Education[2] issued by the United States Bureau of Education, which shows that at present between three-fourths and one per cent of our whole population is enrolled in college, whereas in 1890 only about one-fourth of one per cent was so enrolled. The figures given also show that the last five-year period has exhibited a much greater increase than any other of similar length.

Concurrent with the tendency just stated have been a considerable decline in the purchasing power of the dollar and a general demand that the scope of education be enlarged. The united effect of these three factors has been such that it is practically impossible to secure the amounts of money necessary to provide what are considered adequate educational facilities for all those who wish to enjoy them. The difficulty of doing so appears to be greater in the field of higher education than in any other.

One outstanding question, which has arisen in connection with the crucial situation just described, is that of whether or not institutions of higher education shall open their doors to practically all those who have completed a secondary course and wish to enter. The general tendency has been for state supported institutions to approximate doing

[1] ODELL, C. W. "Are college students a select group?" University of Illinois Bulletin, Vol. 24, No. 36, Bureau of Educational Research Bulletin No. 34. Urbana: University of Illinois, 1927. 45 p.

[2] PHILLIPS, FRANK M. "Statistical summary of education 1923-24." U. S. Bureau of Education Bulletin, 1926, No. 19. Washington, 1926. 7 p.

so, whereas those deriving their support from other sources exercise varying degrees of selection among the applicants for admission. These policies are rarely based upon any thoroughgoing study of the problem and never upon conclusive evidence as to the best practice, so that what should be done may be considered still an open question. The data available[3] appear to warrant the statement that at present the group of those who actually enter college represents a marked selection of all high-school graduates, but that it still contains many individuals who apparently cannot carry the usual type of college work successfully. On the whole, therefore, it seems desirable, perhaps even necessary, that if colleges[4] are to continue to maintain their present scholastic standards, some degree of selection among applicants for admission should be exercised. Such an assumption naturally raises the question as to what is the most desirable basis of making this selection. In other words, do any of the data which are fairly readily obtainable concerning high-school graduates provide a satisfactory, or even a helpful, basis of foretelling scholastic success in college? If so, which of these data are most valuable for this purpose and how much confidence should be placed in their use?

A second question arising, in part at least, from the same cause and one which has attracted much attention recently is that of providing for college students of different aptitudes and abilities. This question really divides itself into two parts. In the first place, if the amount of selection at college entrance is not very great it will undoubtedly result that many of those who are allowed to enter can, or at least will, not do satisfactory work in certain subjects, whereas in others they will do passing or even superior work. The college is therefore confronted with the need for providing educational guidance for such students. This requires, if possible, the determination of the subjects or courses in which the students will succeed and those in which they will fail. Even if entrance requirements are decidedly severe and many of those seeking admission are barred, educational guidance of the sort just mentioned is still desirable though the necessity for it is not so acute. In the second place, there has recently been considerable interest in the matter of offering different types or levels of instruction within single subjects and otherwise varying the educational opportunities given students of different abilities. It is true that this problem has received much more attention in elementary and high schools than in colleges,

[3] Odell, *op. cit.*, p. 26-29.

[4] The term "college" will be used frequently as a general term including all types of institutions of higher learning.

but an increasing number of the latter are giving it serious consideration. In this case also, the less selection there is among applicants for admission to college the greater is the need for attention after admission because the group admitted is more heterogeneous. Even if a relatively high degree of selection is exercised among those who seek to enter college, however, those who gain admission and who enroll for any particular subject practically never constitute a truly homogeneous group. Therefore, there is need to determine the validity of various bases which may be employed for classifying students in advance according to the different amounts and kinds of subject matter, types of instruction, and so forth, which seem best suited to them.

The purpose of this bulletin. It is the purpose of this bulletin to present a study and evaluation of some of the more readily available items of information which may be, and in many cases are, used to predict the probable scholastic success of college students. After reviewing briefly a number of studies illustrative of what has already been done in the field, the writer will give an account of one[5] along this same line which he has been carrying on. This investigation differs from most of the others in the same field in that the attempt has been made not merely to determine the accuracy of prediction of college success in general, but also for each subject carried by any considerable number of the individuals included. It is limited by the fact that the college data upon which it is based include only records for the freshman year. Its purpose may, therefore, be stated as being to show how accurately the marks of college freshmen in their various subjects can be predicted when their ages, scores upon an intelligence test, and complete high-school records are available. The problem will be attacked primarily by the methods of simple and multiple correlation and the accuracy of predictions based on the best multiple regression equations obtainable will be shown.

---

[5]The first part of this study, which included only the data obtained while the individuals embraced were still in high school, has been presented in the following bulletin:

Odell, C. W. "Conservation of intelligence in Illinois high schools." University of Illinois Bulletin, Vol. 22, No. 25, Bureau of Educational Research Bulletin No. 22. Urbana: University of Illinois, 1925. 55 p.

A second portion which deals with the question of how great a selection occurs among college entrants as compared with high-school graduates has been dealt with in the following publication:

Odell, C. W. "Are college students a select group?" University of Illinois Bulletin, Vol. 24, No. 36, Bureau of Educational Research Bulletin No. 34. Urbana: University of Illinois, 1927. 45 pp.

The present bulletin is the third in the series.

CHAPTER II

A BRIEF REVIEW OF WHAT HAS ALREADY BEEN DONE

The extent to which intelligence tests have been used in institutions of higher learning. Since most of the recent studies dealing with the prediction of scholastic success in college have employed intelligence test scores as the chief criterion, it seems in place to mention several studies which show something of the extent to which intelligence tests have been employed in college, both for this and other purposes. Here and later no attempt will be made to refer to all of the investigations which have been reported, but only a few of the most significant or typical ones will be mentioned in each case. Bridges,[1] early in 1922, received answers from 42 of 70 institutions to which he had sent inquiries and found that although 31 of the 42 had made some use of group intelligence tests only a few had done so in connection with determining admission. Apparently, in many cases, the tests were administered with no very definite purpose in mind. A year and a half later Laird and Andrews[2] reported that 26 out of 64 institutions made some use of tests as part of the routine process of determining the admission of applicants and that others used tests for such purposes as sectioning classes, determining the amount of work to be carried, giving vocational and educational guidance, deciding upon the elimination of students, and dealing with disciplinary cases. Probably the most detailed report of the use of intelligence tests in colleges is that by MacPhail,[3] which appeared some three years ago. In this he summarized briefly almost every article dealing with this topic and showed that in many institutions intelligence tests played a definite part in the admission of applicants as well as in other questions of policy. A more recent study by Toops[4] reported that 66 out of 110 institutions answering a questionnaire employed intelligence tests during the year 1923-24. None of these

[1] Bridges, J. W. "The value of intelligence tests in universities," School and Society, 15:295-303, March 18, 1922.

[2] Laird, D. A., and Andrews, A. "The status of mental testing in colleges and universities in the United States," School and Society, 18:594-600, November 17, 1923.

[3] MacPhail, A. H. The Intelligence of College Students. Baltimore: Warwick and York, 1924. 176 p.

[4] Toops, H. A. "The status of university intelligence tests in 1923-24," Journal of Educational Psychology, 17:23-36, 110-24, January, February, 1926.

based admission entirely upon test results, but 19 used them as a partial basis. Forty-nine took them into account in determining dismissal for low scholarship, 34 in determining probation, 36 used the results in determining the amount of work students should carry, 25 in selecting and encouraging bright students to take graduate work, 42 in motivating the work of bright students, and various numbers in selecting assistants, appointing scholars and fellows, and so on.

From the studies referred to above it will be seen that intelligence testing is apparently well established in many institutions of higher learning and that the results receive large use in a number of matters having to do with guidance, instruction, and other direction of students, as well as to a somewhat lesser degree with their admission. So far as the writer knows, no institution has yet based admission upon intelligence test scores alone, though for certain classes of applicants a few colleges make them the chief criterion.[5]

Summary studies of the relationship of intelligence test scores and other criteria to college marks. Several of the studies mentioned and a number of others present data showing the degree of relationship found between college marks and intelligence test scores, high-school marks, and other items of information. Before considering a few reported investigations in greater detail it seems well to give a brief picture of general tendencies. Terman,[6] reporting on 25 colleges, found coefficients of correlation[7] running from .29 to .83 between test scores and college marks, whereas those between the latter and high-school marks ranged from .38 to .74, and those between them and college entrance examination results from .25 to .62. Incidentally, he states that the Thorndike Intelligence Examination is probably the best of those available for the purpose of predicting scholastic success in college. Roberts[8] reports similar ones of .31 to .60, also coefficients between college and high-school marks of .53 to .69 and between the former and college entrance examinations of .25 to .62. He makes this statement, "Combining intelligence scores with all other good measures, the exceedingly high correlations of .75 to .80 are obtained between these measures and

[5]This refers chiefly to the admission or rejection of applicants who have not completed the required secondary school work and who are also above the usual age.

[6]Terman, L. M. "Intelligence tests in colleges and universities," School and Society, 13:481-94, April 23, 1921.

[7]The meaning and interpretation of coefficients of correlation of various sizes is discussed in Chapter VI.

[8]Roberts, A. C. "Objective measures of intelligence in relation to high-school and college administration," Educational Administration and Supervision, 8:530-40, December, 1922.

school marks." He also writes, "The intelligence scores have shown themselves our surest guide in detecting the very highest and the very lowest of intellectual ability." MacPhail[9] lists about 60 correlations between test scores and college marks, ranging from .13 to .71.

The use of intelligence tests at Brown University. Due to the work of Colvin, assisted by MacPhail and others, Brown University has for about ten years been among those institutions making the most extensive and careful use of intelligence tests in connection with the admission of students and also, though perhaps to a lesser degree, in connection with their guidance and direction after entrance. Not only have intelligence tests played a prominent part in determining the admission of freshmen at Brown University, but also a number of articles have appeared describing their use for this purpose. Therefore it seems fitting to select this institution as the example which will be described in more detail than any other as an illustration of what is being done.

The work along this line began during the time of the World War and by 1919 Colvin[10] reported on the first two or three years' use of tests. At this time he stated that different intelligence tests correlated from about .40 to .60 with freshman marks, and that of the students who did unsatisfactory or unusually good work about two-thirds were indicated by the test scores. On the whole the results were considered sufficiently satisfactory to warrant continuing the use of tests. A year later another article[11] by the same writer gives about the same correlations as before, those for the Brown University Psychological Examination and the Thorndike Intelligence Examination being a number of points higher than those for Army Alpha and also being on the whole higher than the corresponding correlations for high-school marks or teachers' estimates. The test results appeared to pick out the superior and inferior students with more accuracy than the average ones. When the Brown and Thorndike scores were averaged 90 per cent of the lowest tenth were found to have failed in one or more subjects.

In 1922 Colvin and MacPhail[12] replied to some unfavorable criticisms of the use of intelligence tests in college and gave further data concerning their use at Brown University. Most of these merely substantiate previous statements, though in some cases they are presented

---

[9] MacPhail, *op. cit.*, p. 29.

[10] Colvin, S. S. "Psychological tests at Brown University," School and Society, 10:27-30, July 5, 1919.

[11] Colvin, S. S. "Validity of psychological tests for college entrance," Educational Review, 60:7-17, June, 1920.

[12] Colvin, S. S., and MacPhail, A. H. "The value of psychological tests at Brown University," School and Society, 16:113-22, July 29, 1922.

in a different form. The writers state, for example, that low test scores furnish a more reliable prediction that college work in general will be poor than do low marks made during the first semester but that a combination of the two is better than either one alone. Of college honors 80 per cent went to those earning high test scores, 19 per cent to those with medium scores, and only 1 per cent to those with low scores.

More recently Burwell and MacPhail[13] have written upon the same topic. They report that the procedure has been changed somewhat by giving the Brown test to all freshmen and the Thorndike test only to the lowest fifth, in place of giving both to all freshmen as had been done for several years. Among the statements made are that "new students who will probably fail in two or more subjects in either semester during their first year in college are far more likely, roughly speaking ten to twenty times more likely, to be found among those who make low psychological scores than among those with high ratings;" that "a freshman whose psychological score places him in the lowest decile has only two chances out of five of remaining more than one year in college and only one chance out of five of graduating;" and, finally, that "the majority of honor men are to be sought among those scoring in the best psychological third; most of the remainder may be expected to come from the middle third; and a very few (about one out of twenty) from the lowest third." Forty-six has been set as a critical score on the Brown University test above which a student must rate to indicate that he will probably receive no grades below "C" during the first semester.

It appears that those who have been using the tests at Brown are very firmly convinced of their value. However, they recognize and point out certain limitations and indicate that it is highly desirable to have other data to supplement the test results, but apparently regard them as the one most important criterion for predicting scholastic success in college.

The use of tests at Columbia University. Columbia University, also, has made rather extensive use of intelligence tests in connection with admitting students. Accounts of the work have been given by Wood,[14] Thorndike,[15] and others. The experiments there appear to have begun in 1919. At that time faculty action was taken providing two possible methods of entrance, one of which was the old method based upon

[13] Burwell, W. R., and MacPhail, A. H. "Some practical results of psychological testing at Brown University," School and Society, 22:48-56, July 11, 1925.

[14] Wood, B. D. Measurement in Higher Education. New York: World Book Company, 1923, Chapters II-V.

[15] Thorndike, E. L. "On the new plan of admitting students at Columbia University," Journal of Educational Research, 4:95-101, September, 1921.

entrance examinations in high-school subjects, previous school records, health records, and estimates of character and personality. The second method substituted intelligence tests for the subject-matter examinations included in the first. For purposes of record all those desiring to enter by the first plan as well as those entering by the second are given the Thorndike Intelligence Examination. Many different sets of figures are given to indicate the validity of this test as used at Columbia for forecasting success in college work. The correlations between test scores and college marks average around .65 and are distinctly higher than those of the latter with college entrance examinations, New York Regents' examinations, and still more so than those with secondary-school marks. The correlations obtained for the test results are probably increased somewhat because no applicant for admission is allowed to take the test unless the data concerning him on the other three points mentioned are satisfactory. The same is, however, true of those admitted with examinations covering high-school subjects as one of the criteria and doubtless raises the correlations there also. In conclusion it may be said that the use of intelligence tests, as one of the bases for determining admission to Columbia University, has become an integral part of the procedure and is no longer considered an experiment.

The use of intelligence tests at the University of Minnesota. The reports[16] from this institution are not as favorable to intelligence tests as those from Brown and Columbia Universities. It appears that high-school marks, the kind of work carried in high school, and marks on three themes at the beginning of the freshman year at the University were all more reliable in indicating students whose university work was poor than were the scores made on a mental test. When the latter were combined with the former, a correlation of about .70 was obtained. It is pointed out that in most cases of marked discrepancy between the work actually done and the predictions made from the combined criteria explanations can be found when the individual cases are studied. What has been accomplished at the University of Minnesota may be summarized as follows: a threshold has been fixed such that only 1 per cent of those falling below it will prove successful in college work; the procedure can be explained to students and all others interested; students of unusual ability can be located; a beginning of vocational se-

[16] JOHNSTON, J. B. "Predicting success in college at the time of entrance," School and Society, 23:82-88, January 16, 1926.

JOHNSTON, J. B. "Predicting success or failure in college at the time of entrance," School and Society, 19:772-76, June 28, 1924; 20:27-32, July 5, 1924.

JOHNSTON, J. B. "Tests for ability before college entrance," School and Society, 15:345-53, April 1, 1922.

lection has been made; promising students not in college can be selected and encouraged to attend; college failures can be treated much more adequately; students who need special advice can be selected and given this advice; and finally each student guided "so far as possible into that line of effort in which his native ability will find its most complete expression."

The use of intelligence tests at other institutions. In view of the fact that there is great similarity between the results reported from most of the institutions which have employed intelligence test scores as one of the criteria for determining admission, it seems not worth while to refer to reports from more than a few different institutions. Those which are mentioned in this section were chosen partly more or less at random and partly because the results obtained were in some way different from the general trend.

The results reported from the University of Pittsburgh[17] are distinctly lower than those given previously. In this case college marks correlated only .41 with Army Alpha Scores, as compared with .32 with first semester marks. These correlations were undoubtedly lowered somewhat by the fact that the individuals included in the study were more highly selected than an ordinary freshman class, and also by the fact that the Army Alpha Test seems, on the whole, not to predict scholarship as well as do the Thorndike, Brown, and several others. Differing much from this is an unusually high correlation reported from the State Normal School at Indiana, Pennsylvania.[18] The National and Illinois Intelligence Tests were used and the scores correlated above .70 with educational psychology mark.

May, at Syracuse University,[19] secured information as to the number of hours spent in study and found that combining this with intelligence test score gave a multiple correlation of .83, whereas test score and high-school mark gave only .64 with honor points in college. He also found that when the amount of study was held constant the correlation between test score and honor points was .81. A study at the University of Washington[20] corroborates this, although it does not present its results in just the same way. Wilson, who reports it, concludes that the

[17]Ernst, J. L. "Psychological tests vs. the first semester's grades as a means of academic prediction," School and Society, 18:419-20, October 6, 1923.

[18]Rich, S. G., and Skinner, C. E. "Intelligence among normal school students," Educational Administration and Supervision, 11:639-44, December, 1925.

[19]May, M. A. "Predicting academic success," Journal of Educational Psychology, 14:429-40, October, 1923.

[20]Wilson, W. R. "Mental tests and college teaching," School and Society, 15:629-35, June 10, 1922.

TABLE I. COEFFICIENTS OF CORRELATION BETWEEN THORNDIKE TEST SCORE AND FIRST SEMESTER COLLEGE MARKS

Biology	.51	History	.46
Chemistry	.43	Human Progress	.69
English	.36	Mathematics	.52
French	.42	Physics	.50
German	.50	Public Speaking	.46
Graphics	.35	Spanish	.57

failure of intelligence tests and college marks to correspond more closely is largely accounted for by the differences in the amounts of time spent in study, especially by the fact that, on the whole, bright students study less than do dull ones.

Studies showing correlation in particular subjects with intelligence test scores and other data. Most of the many studies made have correlated the various criterion measures with college averages, only a few dealing with marks in particular college subjects. Of the few, two which correlated test scores with college marks and one which used high-school marks instead of test scores will be mentioned. One[21] of the first two was made at the University of Pittsburgh and yielded the average coefficients of correlation between score on the Thorndike test and freshman college marks for the first semester shown in Table I. The correlation of the test result with the general freshman average for the first semester was .51. Root, who reports the study, concludes that test results are decidedly valuable for predicting academic success, but that they are only one of the needed items of information. He points out that if the criterion for admission to the university were taken as being the lower limit of the middle group upon the tests, all applicants scoring above that point being admitted and all below rejected, about one-third of the students would be excluded or admitted improperly, that is, excluded when they could do satisfactory work or admitted when they could not.

The other[22] of the two studies does not give tables of the exact coefficients of correlation, but summarizes the results found from correlating Otis test score with college marks as follows: "In all cases the correlations are positive. In all cases on the average the pupils who stand high in the test stand high in scholarship; those who stand low on the test stand low in scholarship, and those who stand in the middle

[21] Root, W. T. "The freshman: Thorndike college entrance tests, first semester grades, Binet tests," Journal of Applied Psychology, 7:77-92, March, 1923.

[22] Jordan, A. M. "Student mortality," School and Society, 22:821-24, December 26, 1925.

TABLE II. COEFFICIENTS OF CORRELATION BETWEEN HIGH-SCHOOL AND COLLEGE MARKS IN CERTAIN SUBJECTS

College Subjects	High-School Subjects						
	Eng.	Chem.	Alg.	Geom.	Lat.	Elem. French.	Adv. French
English	.28	.20	.18	.19	.23	.25	.31
Chemistry	.21	.19	.21	.23	.23	.28	.26
Algebra	.22	.34	.41	.41	.26	.41	.33
Analytic Geom.	.28	.31	.38	.34	.30	.38	.40
Elem. French	.40	.26	.43	.14	.14	...	...
Adv. French	.39	.18	.26	.03	.32	.41	.35
German	.19	.25	.31	.28	.39	.37	.43

on the test are in the middle in scholarship. But in some cases the relationship is quite low, while in other cases it is moderately high. In no case is there a high coefficient of correlation between the test and the marks in any subject. With German during the first year the relationship is quite respectable, but, not even here, high enough for prognosis. The coefficients are quite substantial (from .45 to .61), then, between the Otis test and the marks in German, English, history, geology, and French for the first year; present but low (.32 to .39) in the case of mathematics, chemistry, Spanish, economics, engineering, and Latin. During the second year the coefficients are substantial in English and Spanish; present but low in French, history, economics, engineering, and German; and negligible in mathematics, chemistry, geology, Latin, and zoology. However, the coefficients of correlation during the second year are necessarily lower because of the contraction of the range of scores (the lowest have largely disappeared). The correlations with average and total grades are marked." Jordan also states the correlations obtained between high-school and college marks. These varied from .37 to .59 and on the average were quite similar to the coefficients between college marks and test score. The closest relationships appeared to be in economics, Spanish, and French. Using multiple correlation with combined test score and high-school average mark he obtained a coefficient of .58 with the university average for two years. The conclusion from his study is, therefore, that there is little difference in prognostic power betwen the score on the Otis Group Intelligence Scale, Advanced Examination, and the high-school mark, although the correlations of the former with college marks are lower than those Root found with the Thorndike score. This latter fact is undoubtedly due, at least in part, to the fact that the Thorndike test is considerably longer than that

of Otis and so yields a more satisfactory measure for the purpose here discussed.

The third study referred to[23] was conducted at the University of Maine and dealt with the correlation of high-school and college marks in particular subjects. The correlations found are given in Table II. These coefficients seem to warrant the conclusion that correlations between high-school marks and college freshman marks in single subjects above .40 are rare and that the central tendency of such correlations is not far from .30. This, however, is not supported by the results obtained by Jordan, whose corresponding coefficients averaged about .20 higher.

The other studies of this sort available tend to yield correlations of about .40 to .50 or .55 between test scores and marks in single college subjects and about the same between high-school and college marks. Most of them are based on smaller numbers of cases covering only a few subjects and are hardly worth mentioning separately.

Summary. The work which has been done up to date in attempting to predict the scholastic success of college students by means of intelligence test scores and other criteria may be briefly summarized in the following statements. In a considerable number of institutions of higher learning, including several of the largest ones in this country, the use of intelligence test scores as one of the criteria for admission has passed the experimental stage and is now a settled policy. In many other institutions much use has been made of intelligence tests, but it has had little or no connection with the admission of students. Although the correlations reported vary from near zero up to .70 or above, a range of .40 to .50, or perhaps somewhat higher, may usually be expected between score on an intelligence test and freshmen mark. These correlations are about the same as those of high-school with freshman marks and both slightly higher than those given by entrance examinations covering high-school subjects. The true relationship in the latter case is, however, somewhat closer than indicated by the obtained coefficients of correlation, because applicants for admission making low marks are generally rejected and thus the range decreased and the apparent correlation lowered. If one of the best tests is employed the correlations with freshman mark will probably be higher than will those of the high-school average. A combination of test score and high-school mark may be expected to yield correlations of about .60 or higher.

[23] Gowen, J. W., and Gooch, M. "The mental attainments of college students in relation to previous training," Journal of Educational Psychology, 16:547-68, November, 1925.

If it is possible to include an adequate measure of study habits also the coefficients will probably rise to near .80. Comparatively little has been done in attempting to predict scholastic success in single college subjects, but apparently the correlations found for single subjects are on the whole not much, if any, below those found for the general freshman average and apparently here, also, a test score gives about the same accuracy of prediction as does a high-school mark. Perhaps the most important conclusion is that the whole problem needs much more careful and extensive investigation, especially along the line of finding the best type or types of entrance examinations covering the high-school subjects and of combining the results therefrom with other data to give the best multiple predictions possible.

CHAPTER III

THE GENERAL PLAN OF THIS STUDY

The initial collection of high-school data. The data used in this investigation concern a group of individuals graduated from several hundred high schools in the state of Illinois in 1924 and admitted to various institutions of higher learning in the summer or autumn of the same year. In the fall of 1923 all the four-year public high schools in the state were invited to cooperate with the Bureau of Educational Research in this study. The number that did so was 368, a few more than one-half of all those within the state, and the number of seniors included was about 12,300. The data secured concerning them consisted of their scores upon the Otis Self-Administering Test of Mental Ability, Higher Examination, Form A and the answers to the questions on an "Information Blank for High-School Seniors," which called for the following information:

Name
Sex
Date of Birth
Age on September 1, 1923
Name of school
Town or city
Intentions concerning further education
    Intention of continuing
    Institution
    Course
    Major subject
Vocational choice
Father's occupation
Information as to previous intelligence tests taken
Units of high-school credit
High-school subjects liked most
High-school subjects liked least
Number of failures in high school
Average high-school mark[1]

The tests were given by principals or by teachers designated by them and the information blanks filled out by the seniors themselves. All scoring of test papers and tabulation of results was done in the offices of the Bureau of Educational Research.

[1]This was the average mark up to date or for the first three years. It was secured from only a minority of the schools and for about 2700 seniors.

The second step in collecting high-school data. A year later, in the fall of 1924, the 368 high schools were asked to furnish the complete high-school scholastic records of all pupils for whom the other information had been secured, and also if possible, to state what, if any, institution of higher learning each individual was attending. A few of the seniors of the year before had not been graduated, and in a few cases the desired records were not forthcoming, but the loss from these sources was comparatively slight, so that the complete scholastic high-school records of about 11,500 graduates were secured. Since these marks came from several hundred schools which employed a total of over one hundred different marking systems, if all minor variations be counted, it was necessary to transmute them to a uniform basis. For this purpose a percentile system with passing at 70 and no conditions was chosen. The marks given according to all other plans were changed to this system by approved and careful statistical procedure.

The collection of college freshman data. Some three hundred institutions of higher learning had been named by the seniors in answer to the question as to where they expected to continue their education. Early in the academic year of 1925-26 letters were addressed to all these institutions asking for the complete 1924-25 scholastic records of all freshmen coming from any of the high schools included in this study. About 7,700 of the seniors had stated that they intended to continue their education, in addition to many who were undecided, and the majority of them had named the institutions they expected to attend. Despite this fact the freshman records of not quite two thousand students were all that were secured. This loss is due to at least four causes. In the first place, a number of the collegiate institutions addressed either were unwilling to cooperate in the study, or, after expressing their willingness to do so, failed to send the desired records. A second reason was that a number of the institutions which did cooperate failed to furnish the data for all of their students for whom they were desired. Third, undoubtedly many of the high-school graduates who planned to attend college found it necessary, for financial or other reasons, to postpone entrance for a year or more after high-school graduation. The last, and probably the most important, reason was the fact that in filling out the information blanks the high-school seniors expressed their highest hopes and ambitions or gave answers which they thought would sound best and that, therefore, many of them who had very slight expectations of ever actually attending college, signified that they intended to do so.

Of the approximately two thousand students whose records were secured from various colleges, almost one hundred did not remain in college long enough to have any marks recorded. The number for whom marks for at least one quarter, term, or semester were secured was 1892, and for 1677 of these a full year's marks were obtained. As these marks were given by more than one hundred institutions it was necessary to transmute them to a common basis in the same manner as had been done for the high-school marks, and so all were adjusted to the same basis of a percentile marking system with 70 as passing and no conditions.

The reliability[2] of the data secured in this investigation. There is no doubt that in both intelligence test scores and high-school and college marks large variable errors are present. No group intelligence test so far devised yields highly accurate individual scores and the Otis Self-Administering Test, which requires only half an hour to give, is probably less reliable than one, such as the Thorndike Intelligence Examination, which consumes two or three hours. Moreover, the tests were not administered by a corps of trained and selected examiners, but by several hundred different principals and teachers, many of whom had probably never given a standardized test. This fact undoubtedly served to increase the errors in the scores. It should not be overlooked, however, that the test used reduces the directions to be given by examiners to a minimum and that, therefore, the errors due to lack of training of the persons giving the tests are less than would otherwise be the case. The writer does not believe, however, that this factor of added reliability is sufficient to balance the two of brevity and administration by poorly qualified examiners which make for the opposite effect.

The method of computing intelligence quotients, which Otis provides, introduces a constant error into many of those so determined.[3] However, as little use will be made of the I. Q. in the discussion, it does not seem worth while to discuss this point further than to call attention to the fact that the coefficients of correlation between the I. Q. and other data are probably slightly lower than they should be and,

---

[2]As used in this bulletin, the term "reliability" is practically equivalent to "accuracy." It is not limited to its sometime narrow technical meaning referring to the agreement between two sets of scores on the same measuring instrument, though it includes this.

[3]For a more complete discussion of this point, see:

Odell, C. W. "Are college students a select group?" University of Illinois Bulletin, Vol. 24, No. 36, Bureau of Educational Research Bulletin No. 34. Urbana: University of Illinois, 1927, p. 16-17.

therefore, the estimated accuracy of predictions made on the basis of the I. Q. is also slightly too low.

An additional fact which probably affected the significance of the test scores was that about half of the seniors tested had never taken an intelligence test before and it is likely that many of their scores, when compared with most of those of the seniors who had taken such tests previously, do not fairly represent their mental ability. Furthermore, because of the conditions under which the tests were given, there was generally no particular incentive, apart from the desire to excel, for the pupils to do their best. Hence, it is likely that a considerable number of them did not put forth maximum effort while taking the test. These and all other causes which produce variable or accidental errors in the test scores result in lowering the correlations and other predictive indices based thereon and justify the conclusion that the real relationships are somewhat closer than those actually computed.

Too much evidence and discussion concerning the subjectivity and unreliability of school marks has appeared within the last few years for the subject to need extended comment in this connection. Undoubtedly the errors present in the marks were increased somewhat by the fact that marks from several hundred high schools and more than a hundred colleges with different systems and standards were transmuted to a common basis and thrown into a single group. In spite of the fact that the transmutation was made with great care and followed sound statistical procedure, it was not possible, in all cases, to be sure that the transmuted marks were really equivalent to the original ones. The effect of increasing such variable errors was to lower the coefficients of correlation and other predictive measures secured.

The computation of zero-order coefficients of correlation. As has been suggested, the chief method employed in determining the relationships existing between college freshman marks and the other data available was that of correlation. It was found that there were forty-nine subjects or closely related groups of subjects[4] each of which had been carried by ten or more freshmen. Correlation tables were made for the mark in each of these subjects or subject groups with age, mental test score, intelligence quotient, general high-school average, and average

[4]In a number of cases it is doubtful just what really constitutes a "subject" as the term is commonly used. This, for example, is true of agriculture. In cases in which there were only a few freshmen who carried each of the several possible divisions the procedure followed was to group them together as a single subject. Agriculture, therefore, includes various courses in agronomy, animal husbandry, and so forth; art includes freehand drawing, painting and sculpture, and so on with others.

mark in each high-school subject or group of subjects with which it seemed likely that there was close relationship. Thus, for example, college freshman biology mark was correlated with marks in high-school biology, botany, general science, and zoology and also with the average mark in all high-school science. Likewise, that in French was correlated with high-school English, French, Latin, and Spanish marks, and also with the general high-school foreign language average.

In addition to these correlations quite a number were made between the freshman marks and the amounts of time devoted to particular subjects or groups of subjects in high school and also between freshman marks and those in the work of particular years in high school. For example, the freshman biology mark was correlated with the number of semesters of biology carried in high school, also with the total number of semesters of science carried. The freshman French mark was correlated with the number of semesters of high-school French, of high-school Latin, and of all high-school foreign language, also with the marks in first, second, third, and fourth year Latin and French, in so far as each had been carried. After such correlation tables had been made for a dozen or more of the freshman subjects it appeared that the results therefrom would contribute nothing of value to the study, so no more were constructed. The correlations of freshman marks with the amounts of particular subjects and groups of subjects carried in high school were so near zero as to offer no help in predicting freshman marks. The correlations with a particular year's work in high school were higher, but they appeared to add nothing not already contributed by those of freshman marks with marks in all the high-school work in the various subjects. In some cases they were practically as high as the latter, but the use of the multiple correlation procedure showed that they added almost nothing in accuracy of prediction.

After constructing the tables described, the next step was naturally to compute the simple or ordinary coefficients of correlation for them. It should be remembered that, since many of these correlations involve as one variable an average mark for a group of similar high-school subjects or for all high-school subjects, the coefficients obtained from them are in a sense multiple coefficients although not obtained by the multiple correlation method. In other words, they show the relationship existing between college freshman marks and combinations of several different high-school marks.

The computation of coefficients of multiple correlation and regression. The calculation of zero order or simple coefficients of corre-

lation was followed by that of multiple coefficients and regression equations. In view of the considerable amount of labor involved in computing the latter they were not found for all freshman subjects, but for only about one-third of them. These were in general the subjects carried by the largest numbers of freshmen and two or three others included because of especial interest in them. In connection with this the admission should be made that since many possible correlations were not computed it is probable that some were omitted which should have been found. Since the amount of money available for clerical help, though fairly generous, was not unlimited, it was necessary that the line be drawn somewhere, and it is very likely that the writer's judgment in selecting the most promising possibilities was not infallible. In the case of several of the freshman subjects two or three groupings were made according to the high-school subjects carried and a different set of multiple correlations computed for each grouping. For example, in addition to calculating the correlations and regressions for all freshmen who carried Latin as a college freshman subject, they were also found separately for the portion of this group that had carried high-school French. The general procedure in computing the multiple coefficients was to start with the highest one of zero order and combine the others of the same order with it until the addition of another criterion no longer increased the obtained coefficient by as much as .01. Because of the fact referred to above, that many of the simple coefficients of correlation were really multiple in nature though not in derivation, it could not be expected that on the whole there would be as great an increase in the multiple coefficients over those of zero order as would otherwise have been the case.

The question may be raised as to why certain combinations, which will appear later in the chapter containing the multiple correlation results, were made, in view of the fact that one of the simple correlations already used was that of the freshman mark with the general high-school average or the average in a group of similar subjects, and another the correlation with one of the subjects which entered into this group. For example, the highest obtained multiple coefficient for freshman rhetoric was that obtained from a combination of high-school average, high-school English mark, and point score on the test, and of course the high-school average included the high-school English mark. The reason for so doing is, however, clear to any one familiar with multiple correlation. In computing the high-school general average or the average in any group of similar subjects the marks entering into the given average were all allowed the same weight in determining it. By means of mul-

tiple regression equations, however, one is able to determine the optimum weight which should be given to each factor, that is, the weight to give it so that the highest correlation or predictive power will be obtained. Therefore, the fact that a combination of high-school English mark and the high-school general average resulted in a higher correlation with freshman rhetoric than did the former alone, merely means that the weighting of English equally with other subjects in computing the general average is not high enough to yield the best prediction and, therefore if it is only given equal weighting with the other subjects in this average, it should be introduced again with the relative weight indicated by the multiple regression coefficient to accomplish this purpose.

The direct method of securing the same result would be to use no averages of marks in different high-school subjects, but to consider each as a separate variable or criterion in the multiple correlation work. The reason this was not done was that it would have increased very greatly the amount of calculation necessary without yielding more helpful results than the method used. It would, of course, have shown exactly just which of the subjects entering into the high-school average were useful for making the best prediction in each case and which were not. but there seems little advantage in knowing this, provided one knows how to make as good an estimate without this knowledge and with even less labor. Not only was much work saved in computation, but also in the use of results, since the multiple coefficients and regression equations secured involve, on the whole, fewer variables or criteria than would be the case if averages of high-school subjects had not been taken and therefore require less computation in employing them for predictive purposes. The objection can be raised that there are included in the general high-school average marks made in subjects which show much lower correlations with the freshman subject being considered than do those of certain other high-school subjects and that the inclusion of these marks may have lowered the correlation between the freshman subject mark and the high-school average. This contention is true, but the writer believes that for all practical purposes any such results have been taken care of by including in the multiple correlations and regressions the subjects which appeared at all likely to make any contribution to them. Thus, for example, if freshman French mark was best predicted by a combination of high-school marks in English, French and Latin and point score, rather than by including the general high-school average, the method of computation used eliminated the latter. In any event, in view of the practical limitations of time and money, it seemed wise, if not absolutely necessary, to follow the method described above.

The measures of accuracy of prediction obtained in this study. Finally, as a measure of the accuracy or reliability of predictions based upon coefficients of correlation and regression equations, the coefficients of alienation and the probable errors of estimate corresponding to each of the former expressions were determined. The first of these,[5] the coefficient of alienation, is an expression which shows the relationship between the prediction based upon a given coefficient of correlation and a pure guess. For example, the coefficient of alienation which corresponds to a correlation coefficient of .65 is approximately .76. This means that if two variables or series of scores correlate .65 with each other, the estimates of particular scores in one series based upon corresponding known scores in the other will on the average be in error by about .76 as much as if the errors resulted from pure guesses, or, subtracting .76 from 1.00, that the errors will be .24 smaller than those in pure guesses.

The probable error of estimate describes the same situation by stating the limits within which half of the errors will fall. For example, if the probable error of estimate is found to be 4 points on a percentile scale, it means that half of the estimated scores will not vary from the true scores by more than 4 per cent, and, of course, that the other half will differ by more than this amount. These two indices, the coefficient of alienation and the probable error of measurement, give a more concrete and meaningful description of the accuracy of prediction than does the coefficient of correlation.

---

[5]For a more complete discussion of the coefficient of alienation and the probable error of estimate, see Chapter VI. Also:

Odell, C. W. "The interpretation of the probable error and the coefficient of correlation." University of Illinois Bulletin, Vol. 23, No. 52, Bureau of Educational Research Bulletin No. 32. Urbana: University of Illinois, 1926, p. 28-32 and 41-45, and

Odell, C. W. Educational Statistics. New York: The Century Company, 1925, p. 173-74, 230-41, or some other text on the same subject.

CHAPTER IV

THE SIMPLE CORRELATIONS BETWEEN FRESHMAN MARKS AND THE OTHER DATA COLLECTED

The simple correlations computed in this study. At the risk of repeating a portion of the outline of the study given in the last chapter, it seems worth while to state again what correlations were and were not found. The simple or zero-order coefficients obtained are shown in Table III, the first column of which gives the correlations of the freshman marks with age, the second those with point score, the third with I. Q., and the fourth with the general high-school average. Following this are the coefficients found between the marks in various freshman subjects and those in high-school subjects or groups of subjects selected as being most similar to the freshman ones, or as most likely to exhibit significant correlations with them. Thus, for example, the first row of the table shows that freshman accountancy mark had a correlation of —.18 with age, .28 with point score, .29 with I. Q., .47 with high-school average, .38 with high-school commercial average and .47 with high-school mathematics average. As was mentioned in Chapter III, correlation coefficients between certain possible criteria and college marks are not included in this table because, after computing quite a number of them, it appeared that they were of so little value for the purpose of this investigation as not to be worth further consideration. These were the coefficients of the freshman subject marks with the amounts of particular subjects carried in high school and with particular years' marks in high-school subjects, rather than with the average for all of each subject. It will be noted that a number of the coefficients given in Table III are enclosed in parentheses. These are the ones which, because of the joint effect[1] of their small size and the few cases concerned, are less than twice their standard errors or three times their probable errors and so can hardly be considered reliable. The chances are greater than twenty-one or twenty-two to one that all of the co-

[1]The formula for the standard error of a coefficient of correlation is $\frac{1-r^2}{\sqrt{N}}$ and that for the probable error .6745 $\frac{1-r^2}{\sqrt{N}}$, in which r is the coefficient of correlation and N the number of cases. Thus the greater the coefficient and also the greater the number of cases the smaller is the error and the greater the reliability of the coefficient.

CORRELATION OF FRESHMAN MARKS WITH AGES, POINT SCORES, INTELLIGENCE Q

HIGH-SCHOOL SUBJECTS AND GROUPS OF S

Biol-ogy	Bot-any	Chem-istry	Civ-ics	Com-merce	Econ-omics	Eng-lish	For-eign Lang.	French	Gen-eral Sci.	Geo-gra-phy	Geom-etry	G m
				.38					(−.05)			
						.36						
.37	.33								.33			
	(.31)								.30			
		.50							.35			
				.50								
			.23		.28							
						.40	.50	.46			.41	
									.31	.50		
									(.21)			
											.34	
						.46	.49					
							(.39)					
							(.13)					
						.56	.39	.50				
						(.17)						
		.44										
						.65						
		(−.04)							(.14)		.33	
									.53			
			.26		(.35)							
						.39						
						.47						
			.49									
						.49	.52	.60				
				.61								
											.48	
.35	.87								.25			

g to each, they are not equal to more than twice their standard errors, or three times their probable errors. Those without p
e averages of the columns above them.

TS

is- ry	Home Econ-omics	Latin	Man-ual Train.	Math-emat-ics	Mech. Draw-ing	Music	Phys-ics	Pub-lic Speak.	Sci-ence	Span-ish	Zool-ogy
				.47					.46		
				.48							
				.52							
						(−.01)					
									.36		.41
									.47		
	(−.13)						.31		.46		
36									.58		
			(.27)	.47							
		.47								.42	
				.44					.37		
									.29		
				.34							
		.41									
43	.50										
									.60		
			.26								
		.29								(.35)	
				.27	.23						
						(.03)					
						.30					
									.42		
32)				.29			(−.24)		(.22)		
25									.42		(.43)
									.31		
		.40						(.40)			
29		.45								.40	
				.52							
									.47		

re all larger than this ratio.

Freshman Subjects	Age	Point Score	I. Q.	Aver-age	Agri-cul-ture	Alge-bra	Arith-metic	Art
Accountancy	−.18	.28	.29	.47				
Agriculture	(−.09)[1]	.30	(.25)	.51	.59			
Algebra	−.17	.32	.31	.46		.45		
Arithmetic	−.24	.46	.46	.40		.47	.30	
Art	(.03)	.25	.23	(.08)				.44
Ath. Coaching	(.11)	(−.19)	(.22)	(.07)				
Bible	(.02)	(.12)	(.12)	.37				
Biology	(.02)	.20	.20	.43				
Botany	−.32	.44	.42	.51				
Chemistry	−.12	.33	.30	.43				
Clothing	(.05)	(.21)	(.24)	(.10)				
Commerce	−.17	.42	.46	.34				
Dentistry	(−.16)	(.24)	(.23)	.69				
Economics	−.11	.28	.28	.38				
Education	−.23	.32	.32	.37				
Engineering	(.15)	.43	.46	.36				
French	−.24	.32	.32	.51				
General Math	−.28	(.11)	(.14)	.48		.33		
Geography	−.20	.40	.39	.41				
Geology	(.09)	(.22)	(.16)	(.21)				
Geometry	(.03)	(.13)	(.13)	.31				
German	−.24	.27	.28	.42				
Greek	(−.26)	(.10)	(.04)	(.40)				
History	−.12	.27	.28	.43				
Home Econ	(.00)	(.22)	(.30)	(.21)				
Horticulture	(−.10)	(.34)	(.36)	.62	(.31)			
Hygiene	−.24	.22	.23	.33				
Indus. Arts	(−.17)	.25	.24	.29				
Italian	(−.09)	(.01)	(.12)	(−.01)				
Latin	−.24	.39	.41	.51				
Library	−.25	.28	.29	.35				
Mech. Drawing	(−.05)	.16	.16	.19				
Military	(−.03)	(.04)	(.05)	.17				
Music—Performance	(−.05)	(.11)	(.12)	.39				
Music—Theory	−.23	.43	.44	.38				
Pharmacy	(−.30)	.53	.52	.39				
Philosophy	−.44	(.01)	(.12)	.58				
Physical Educ	(−.02)	.18	.16	.15				
Physics	(−.14)	(.19)	(.14)	(−.17)		(.20)		
Physiology	−.36	.46	.47	.53				
Pol. Science	(.02)	.33	.33	.26				
Psychology	−.19	.41	.43	.44				
Pub. Speaking	(−.08)	.25	.27	.40				
Rhetoric	−.21	.35	.38	.47				
Sociology	(−.11)	(.25)	(.23)	.31				
Spanish	−.25	.27	.27	.51				
Stenography	(−.13)	(.29)	(.31)	.48				
Trigonometry	(−.08)	.28	.29	.47		.47		
Zoology	−.20	.38	.37	.58				
Freshman Aver.[2]	−.23	.38	.38	.55				

[1]The coefficients enclosed in parentheses are not reliable. Because of the small number of cases contributin
[2]The coefficients following "Freshman Aver." are those of the average freshman mark in all subjects, not th

efficients not in parentheses are significant or reliable and for most of them the chances are very much greater than this.

The correlations between freshman marks and age. A glance at the first column of the table shows, that, as one would expect, most of the correlations with age are negative. In fact none of the few small positive ones are reliable and the smallest reliable negative one is —.11 for economics. Others close to this are those for chemistry and history. From this point they range up to —.44 for philosophy, the only others greater than —.30 being for botany and physiology. The correlation of the general freshman average with age is —.23. Although about half the coefficients in this column are reliable and therefore indicate that there is a definite inverse relationship between age and freshman marks, they are so small as to offer practically no assistance in predicting the quality of freshman work when age alone is known. Later, in Chapter VI, the question of just how much relationship is indicated by coefficients of correlation of given sizes will be discussed and thus the meaning of these and the others obtained in this study made more concrete. For the present it is sufficient to say that the only prediction justified upon the basis of age is that there is a very slight tendency for freshmen who are below the average age of their group to do better work than is done by those above the average age.

The correlations of freshman marks with point scores and intelligence quotients. An inspection of columns two and three of the table reveals what anyone familiar with the situation would anticipate, that in most cases the entries in the two columns are very nearly the same, the only exceptions being in cases where the coefficients are too small to be reliable. In other words, because of the fact that the point score is one of the two factors upon which the I. Q. directly depends, the correlations of any other variable except age, which is the second factor, with these two are very likely to be the same or almost the same. The writer seriously considered the advisability of not computing any correlations with the I. Q., but did so to try to determine whether, on the whole, it makes any difference at all which one of the two is used. By looking at the columns it will be seen that sometimes one and sometimes the other is the larger and that the coefficients of both with the general freshman average are .38. Thus it appears that from the standpoint of prediction it makes no difference which one is used. From the practical standpoint, however, it seems clear that the point score should be used since the calculation of the I. Q. involves an additional step.

It will be seen that all of the correlations between freshman marks and test results are positive except in the case of athletic coaching and

that here they are not reliable. The smallest ones possessing reliability are those for mechanical drawing and physical education, which are about .16. Other rather low ones at or below .25 are those for art, biology, hygiene, and industrial arts. In addition to these a number of those in parentheses, in fact almost all of them, are small. The closest relationship between freshman marks and test results appears to be in the case of pharmacy for which the coefficients are slightly above .50. Other subjects with coefficients above .40 are arithmetic, botany, commerce, engineering, music theory, physiology, and psychology.

Table III and the discussion just above indicate that there is a closer relationship between score on the test used and freshman marks than between age and such marks. It may be said that the degree of relationship is about twice as close in the sense that it is twice as far away from zero correlation or no relationship at all. A test score, therefore, offers a better basis of predicting success in freshman college work than does an age, but, as will be shown in more detail in Chapter VI, it cannot be said to be very satisfactory for that purpose. Even in the case of pharmacy, which exhibited the closest relationship, estimates of freshman marks based upon test scores would be only 15 per cent better than pure guesses, whereas for the freshman average they would be only about 7½ per cent better. The corresponding figures for age are about 10 per cent better for philosophy, which has the largest negative coefficient, and less than 3 per cent better for the freshman average. Thus the most favorable statement which can be made about predicting freshman marks from the mental test scores secured in this study is only slightly stronger than that made above concerning the use of age for the same purpose. There is a positive, but not very strong, tendency for those who made high scores on the test also to make high freshman marks.

The correlations of freshman marks with the high-school average. The entries in the fourth column of the table, which are the coefficients between the freshman marks and the general high-school average, are, on the whole, higher than those in the preceding columns. Two of those in parentheses are negative and the smallest reliable one, .15 for physical education, is just about the same as the smallest for point score or I. Q. Others below .20 are those for mechanical drawing and military. On the other hand, however, there are a number of these coefficients greater than .50 and two above .60 as compared with only one for point score and I. Q. above .50. The two referred to as being above .60 are .69 for dentistry and .62 for horticulture. Others between .50 and .60 were found in the cases of agriculture, botany, French, Latin, philosophy, physiology, Spanish, and zoology. The correlation between the high-

school and the freshman averages is .55, almost half again as large as that of the latter with test results.

Despite the decided increase in the size of the coefficients, the accuracy of prediction based upon high-school averages is still not at all high. For a coefficient of .69, such as is possessed by dentistry, a prediction is only about 28 per cent better than a pure guess and for one of .55, such as that for the freshman average, it is about 16 per cent better. In other words, just as the improvement over no predictive power at all is roughly twice as great for the test scores as it is for ages, so also it is about twice as great for high-school averages as for test scores. There is no question but that with a few exceptions much better predictions of probable freshman marks could be based upon the high-school averages than upon the test scores obtained in this study. The most marked exceptions to this were in the case of art, commerce, and pharmacy, although in arithmetic, engineering, music theory, and physical education the correlations of freshman marks with the test score were also slightly higher than those with the high-school average.

The correlations between freshman marks and those in single high-school subjects or groups of subjects. The part of Table III to the right of the column headed "Average" shows the coefficients between the freshman subject marks and those in particular high-school subjects or groups of subjects. The number for the different freshman subjects varies from one to eight, but in no case are there more than five of the coefficients for a single subject which are significant. Glancing over them one sees that with the exception of two or three of those in parentheses none are negative and that the reliable ones range from .23 up to .87. The one of .87, between freshman zoology and high-school botany mark, is based on a small number of cases and although it is much more than three times its probable error, probably should not be considered as having high reliability. The next in size is one of .65 between freshman philosophy and high-school English mark which, although not based on a very large number of cases, can still probably be considered fairly reliable. The only others as high as .60 are those of freshman horticulture mark with that in all high-school science, freshman Spanish with high-school French and freshman stenography with high-school commercial work, of which only the second is based on a large number of cases, although the other two are more than three times their probable errors. It will be seen that the central tendency of these coefficients is around .40, about two-thirds of them being between .30 and .50.

In view of the irregularity in the size of the coefficients, it is rather difficult to say that many high-school subjects or groups of subjects are on the whole of more value in predicting freshman marks because they correlate more closely with them, than are others. French, however, may be pointed out as probably the most useful in this connection, since it correlates with freshman French almost as highly as does the general foreign language average, with freshman Latin almost as highly as does English, and with freshman Spanish eight points more highly than any other subject or group of subjects. Thus on the whole it furnishes a better prediction of success in Latin and the two languages derived from it than does Latin itself or the whole high-school foreign language average. High-school English, also, on the whole correlates fairly closely with a number of freshman subjects, though in some cases the coefficients fall below .40.

Looking through the table carefully one will see that in the case of almost every freshman subject the correlation with some one high-school subject or group of subjects is higher than that with the high-school average and also with age, point score or I. Q. In most cases, however, it is not much greater than that with the high-school average. This difference, however, warrants the conclusion that if simple correlation alone is to be used in predicting success in freshman subjects, it will in most cases be unnecessary to secure age records, scores on the mental test used in this study, or even the general high-school average, but that the high-school mark in the subject or group of subjects most similar to the freshman subject is usually the best criterion. As was suggested in Chapter III, however, accuracy of prediction can usually, if not always, be increased by using multiple rather than simple correlation, or, in other words, by basing prediction upon two or more of the items of information rather than upon a single one. How much doing so increases accuracy of prediction will be shown in the next chapter.

Summary and comparison of the results secured in this study with those in the studies of others. On the whole the degrees of relationship of freshman marks with age, test score and high-school marks found in this study are not very different from those obtained by other investigators. Those with age are so small as to be negligible for purposes of prediction. The correlation between freshman average and test score, .38, is slightly, but not a great deal, lower than the central tendency of a large number of studies. Undoubtedly this is accounted for by two facts referred to in discussing the reliability of the results in Chapter III. These are that the test used is considerably shorter than

most tests employed for the same purpose and therefore does not yield as reliable measures and that marks from a great many institutions were grouped together, thus introducing more variations in standards than would be found in the marks of a single institution. The correlations between freshman and high-school subjects found by the writer tend to run about the same as those obtained in the two studies of this sort to which reference is made.

Combining the evidence from all the studies along this line with which the writer is familiar the statement seems warranted that in general a score on any one of the best intelligence tests and the proper high-school mark have about equal value in predicting probable freshman marks. In each case the general expectation concerning the size of the coefficient of correlation is that it will be somewhere between .40 and .50, though under the best conditions one can reasonably expect to secure at least some simple correlations of .60 or higher.

CHAPTER V

THE MULTIPLE CORRELATIONS BETWEEN FRESHMAN MARKS AND THE OTHER DATA COLLECTED

The multiple coefficients of correlation computed. Although a brief statement as to what multiple correlations were found was made in Chapter III it seems best to repeat it here in somewhat more complete form. Seventeen subjects from among the 49 carried by ten or more freshmen each were selected for this procedure. All but one or two of these were the subjects carried by the largest numbers of freshmen, these one or two being added because of some especial interest in them. In three of the subjects, chemistry, French, and Latin, two sets of multiple correlations were computed, one for all, or almost all, freshmen carrying the subject and another only for those who had also carried certain high-school subjects. In Spanish, three sets of coefficients were found, two special groupings being made according to the high-school subjects carried. Thus, including the general freshman average, 23 sets of multiple coefficients were computed.

The procedure in computing the multiple coefficients was first to select all of the simple coefficients of correlation which seemed worth using, the number so selected varying from two to six, and then to combine these to secure the multiple ones. The two always used were the general high-school average and the point score. In addition to these the high-school average in the group of subjects and the mark in the one or more single subjects most similar to the college freshman subject were also included, except in two or three cases in which no high-school subjects that could be said to be similar had been carried by enough pupils to be worth including. The one marked example of this was physical education, almost none of the freshmen who carried this subject having marks recorded for any similar work in high school. In a few cases age was used. In a number of instances the computations were begun with more criteria than were carried through to the finish, since, as the work progressed it could be seen that some of those used made no contributions. In each case the work with those included was carried to the point that no further increase as great as .01 was obtained by computing multiple coefficients of higher orders. In the majority of cases the number of criteria required to accomplish this end was three, in three cases four were required and in none more; in one case the highest zero order coefficient could not be increased and in the

remaining ones two were all that was necessary. In this connection the reader should again be reminded that many of the criteria used, such as the general high-school average, the high-school science average, the high-school mathematics average, and the high-school foreign language average, were themselves combinations of marks in several different subjects and therefore the simple correlations with them were in a sense multiple, although not computed by multiple methods. Because of this fact the increases above the simple coefficienets were not nearly as great as if no such averages had been used.

The multiple coefficients of correlation obtained in this study. The highest multiple coefficients obtained in this study along with related data are presented in Table IV. The first set of four columns therein is for the highest simple coefficient of correlation obtained for each of the subjects mentioned, the second group of four for the highest multiple coefficient and the last group of three for the increase of the multiple over the simple coefficient. Within each of the two groups of four the first column, headed "r" and "R," contains the actual coefficients of correlation, the second, headed "k," the corresponding coefficients of alienation, the third, headed "$P.E._{est.}$," the corresponding probable errors of estimate and the fourth the one or more criteria[1] used in the correlations. The last three columns contain, in order, the increases in the highest multiple over the highest simple coefficients of correlation and the accompanying decreases in the coefficients of alienation and the probable errors of estimate.[2] For example, taking the first line of the table, the highest simple coefficient of correlation of freshman algebra mark with any single criterion was .52, the corresponding coefficient of alienation was .85, the probable error of estimate 6.4 and the criterion, high-school mathematics average. The highest multiple correlation obtained for algebra was .53, with a coefficient of alienation of .85 and a probable error of estimate of 6.4. It was based on two criteria, high-school mathematics average and high-school general average. The increase in the coefficient of correlation was .01, whereas there was no change in the coefficient of alienation or the probable error of estimate.

[1]It will be noted that in a few cases in the table the abbreviations for two of the criteria are connected by the word "or." This means that in such cases the correlation based on the two was the same or so nearly the same that it makes no appreciable difference which one is used. For example, in the case of geometry, the simple correlations with high-school geometry mark and high-school mathematics mark differed by only .0004, so that it makes no material difference which one is used.

[2]It should be remembered that an increase in the coefficient of correlation and decreases in the coefficient of alienation and the probable error of estimate indicate closer relationship or greater accuracy of prediction.

TABLE IV. HIGHEST SIMPLE AND MULTIPLE COEFFICIENTS OF CORRELATION BETWEEN FRESHMAN MARKS AND VARIOUS CRITERIA USED IN PREDICTING THEM

College Freshman Subject	Highest simple correlation with any one criterion				Highest multiple correlation with several criteria				Increase[1] of highest multiple over highest simple correlation		
	r^2	k	$P.E._{est}$	Criterion[3]	R	k	$P.E._{est}$	Criteria[3]	R−r	k	$P.E._{est}$
Algebra	.52	.85	6.4	Math.	.53	.85	6.4	Math., Aver.	.01	.00	.0
Chem. I[4]	.49	.87	6.6	Chem.	.53	.85	6.5	Chem., Sci., P. S.	.04	.02	.1
Chem. II	.46	.89	6.3	Sci.	.49	.87	6.2	Sci., P. S.	.03	.02	.1
Economics	.39	.92	5.3	Aver.	.43	.90	5.2	Aver., Hist., P. S.	.04	.02	.1
Education	.36	.93	4.2	Aver.	.42	.91	4.1	Aver., P. S., Age	.06	.02	.1
French I[5]	.55	.84	4.5	Aver.	.60	.80	4.3	Aver., Fr., Eng., P. S.	.05	.04	.2
French II	.55	.84	5.4	Aver.	.57	.82	5.3	Aver., Eng., P. S.	.02	.02	.1
Gen. Math.	.50	.87	6.0	Aver.	.53	.85	5.9	Aver., P. S., Age	.03	.02	.1
Geometry	.34	.94	7.3	Geom. or Math.	.35	.94	7.3	Geom., Math., Aver. or P. S.	.01	.00	.0
German	.46	.89	5.9	Eng.	.50	.87	5.8	Eng., P S., For. L.	.04	.02	.1
History	.43	.90	5.4	Aver.	.46	.89	5.3	Aver., P. S., Hist.	.03	.01	.1
Hygiene	.32	.95	4.9	Aver.	.34	.94	4.9	Aver., Age	.02	.01	.0

TABLE IV. Concluded

College Freshman Subject	Highest simple correlation with any one criterion				Highest multiple correlation with several criteria				Increase[1] of highest multiple over highest simple correlation		
	r^2	k	$P.E._{est}$	Criterion[3]	R	k	$P.E._{est}$	Criteria[3]	R−r	k	$P.E._{est}$
Latin I[6]	.58	.81	4.9	Aver.	.61	.79	4.8	Aver., Eng., Fr., P. S.	.03	.02	.1
Latin II	.56	.83	4.7	Eng.	.63	.78	4.4	Aver., For. L., P. S., Eng. or Lat.	.07	.05	.3
Phys. Ed.	.18	.98	5.7	P. S.	.20	.98	5.7	P. S. Aver.	.02	.00	.0
Psych.	.43	.90	4.7	Aver.	.50	.87	4.5	P. S., Aver.	.07	.03	.2
Pub. Sp.	.40	.92	5.0	Aver.	.43	.90	4.9	Aver., Eng., P. S.	.03	.02	.1
Rhetoric	.47	.88	4.8	Eng. or Aver.	.52	.85	4.6	Eng., P. S., Aver. or Lat.	.05	.03	.2
Span. I[7]	.51	.86	5.5	For. L.	.56	.83	5.3	Eng., P. S., For. L.	.05	.03	.2
Span. II	.61	.79	5.1	Aver.	.63	.78	5.0	Aver., Span., For. L.	.02	.01	.1
Span. III	.63	.78	5.3	For. L.	.63	.78	5.3	For. L.	.00	.00	.0
Trig.	.51	.86	6.4	Math.	.52	.85	6.3	Math., Aver., P. S.	.01	.01	.1
Average	.55	.84	3.8	Aver.	.58	.81	3.7	Aver., P. S.	.03	.03	.1

[1]The entries in the first column under this heading are actual increases in the coefficients of correlation and those in the other two, decreases in the coefficients of alienation and probable errors of estimate. In all cases, however, they indicate the increases in accuracy or reliability of prediction.

[2]The values of "r" given in this column are not in all cases the same as those given in Table III. This is due to the fact that those given here are the highest of the ones available for multiple correlation and in some cases do not include quite all of the cases which contribute to those in Table III.

[3]The abbreviations in these columns are of the different high-school subjects and groups of subjects, except "P.S." which is used for point score.

[4]Chemistry I includes only those freshmen who had carried high-school chemistry, whereas Chemistry II includes all carrying the subject as freshmen.

[5]French I includes only those who had carried both Latin and French in high school, whereas French II includes all who had carried Latin.

[6]Latin I includes all those who had carried high-school French, and Latin II all who had carried high-school Latin.

[7]Spanish I includes all who had carried high-school Latin, Spanish II all who had carried high-school Spanish, and Spanish III all who had carried high-school French.

The increases of multiple over simple indices of relationship. A comparison of the multiple with the simple coefficients shows that on the whole there was little increase in the latter. In one case out of the 23 the highest simple coefficient could not be raised by including other criteria. The median increase produced was only .03 and the greatest .07, with corresponding decreases in the coefficient of alienation ranging from zero to .05, over three-fourths of them being .02 or smaller. As regards the probable error of estimate over half of the cases were decreased by .1 and in only one case was the difference more than .2. In other words, the increased reliability of prediction obtained in this study by using the best multiple correlations and regressions is so small that it is very doubtful if it can be said to be worth the additional labor and expense required.

The criteria of highest predictive value. In addition to the fact just mentioned probably the one of chief interest in the table is the question of what criteria are most valuable as bases for predicting freshman marks. It was shown in the preceding chapter that in many cases the high-school average yielded the highest correlation with the freshman mark, whereas in many others some similar subject or group of subjects did so. Of the 25[3] criteria employed in the simple correlations given in this table, the high-school average appears in 13 cases, the average mark in a similar group of high-school subjects in six, that in a similar high-school subject in five, and the point score only a single time, for physical education. The criteria used in obtaining the multiple coefficients run very similarly except that the point score appears a much larger number of times. In approximately three-fourths of the cases the high-school average is one of the criteria and the same is true of the point score. Single subject marks occur somewhat less frequently and those in groups of subjects in less than half of the cases. It appears that although the simple correlations between the point score and freshman marks are in general decidedly lower than those of the latter with the high-school average and with marks in various subjects and groups of subjects, yet the point score makes a contribution in prediction somewhat distinct from that made by the other criteria mentioned.

Summary of this chapter. Multiple coefficients of correlation were computed for about one-third of the freshman subjects, in a few cases more than one set being computed for each subject. These coefficients run from .20 up to .63, most of them being between .40 and .60, al-

[3]This number includes the two criteria giving practically equal coefficients in geometry and rhetoric.

though five are at or above the latter figure. The accompanying coefficients of alienation range from .98 down to .78, most of them being in the eighties and the probable errors of estimate from 7.3 down to about half that amount. For the general freshman average the multiple coefficient is .58, the coefficient of alienation .81 and the probable error of estimate 3.7. In no case were more than four criteria needed to secure the highest coefficient and in most cases only two or three. The increases in the multiple above the simple coefficients and the decreases in the corresponding coefficients of alienation and probable errors of estimate are so small that the very slightly increased reliability of prediction appears not to be worth the additional labor of computing. With two or three bare exceptions the estimates possible are still four-fifths or more pure guesses and in the case of about one-fourth of the subjects they are nine-tenths or more pure guesses.

CHAPTER VI

THE ACCURACY OF PREDICTIONS BASED UPON THE OBTAINED COEFFICIENTS OF CORRELATIONS

Purpose of this chapter. To anyone who has not had considerable experience in dealing with predictions when the degree of relationship is expressed by coefficients of correlation the mere statement that a coefficient of correlation is so much, .35 or .60 for example, generally has little definite meaning, especially as indicating how accurate the predictions are. In the two chapters dealing with the correlations found in this study a few brief references have been made to their interpretation in other terms, but it seemed best to devote a chapter to a more elaborate treatment of the matter. The attempt will be made to show in two or three ways just how many and how great are the errors present in predictions associated with coefficients of correlation of sizes typical of those obtained in this and other similar investigations. The writer has discussed the matter at somewhat greater length elsewhere[1] and will not attempt to reproduce in full what has been said there, but will repeat a portion of it with some additional suggestions concerning interpretation.

Interpretation of the coefficient of correlation in terms of the coefficient of alienation or of a "pure guess." In connection with the coefficients of correlation presented in Chapters IV and V some figures were given and statements made as to what they meant in terms of pure guesses. For example, in one place it was stated that a coefficient of .52 indicated that predictions based thereon were only 15 per cent better than pure guesses. To understand this and similar statements one needs to know what is meant by a pure guess. It assumes that the person making the prediction or guess knows what the distribution of the measures which are being predicted is, but does not have any information at all which helps him to determine which measure belongs to any particular case. For example, if one were making a pure guess concerning the marks to be assigned the members of a high-school freshman algebra class from their previous records it would be assumed that he knew the distribution of marks which would be given, but that he had no information at all as to which mark would be given to each individual pupil.

[1] Odell, C. W. "The interpretation of the probable error and the coefficient of correlation." University of Illinois Bulletin, Vol. 23, No. 52, Bureau of Educational Research Bulletin No. 32. Urbana: University of Illinois, 1926, p. 28-32 and 39-45.

The conception of a pure guess can probably be made more meaningful by employing a concrete example. For this purpose, let us suppose that in a high-school freshman algebra class the teacher has decided that the marks she will issue will consist of two A's, seven B's, twelve C's, five D's and four E's. Let us suppose further that someone is told that this distribution of marks is to be given and that, without knowing anything about the individual pupils which in any way concerns their scholastic ability or achievement, he attempts to predict the marks each one will receive. He will, of course, select two individuals as those who will receive A's, seven as those who will receive B's, twelve C's and so on. Since his selections or predictions are not based upon any knowledge whatsoever which helps him in determining the marks assigned individual students, they will be subject to the same errors as if any purely chance method was used, such as placing the names of the pupils in a hat and predicting that the first two drawn would receive A's, the next seven B's, and so on. In either case the predictions made are pure guesses and the relationship between them and the actual marks is represented by a coefficient of correlation of zero.[2]

For predictions to be a certain per cent better than pure guesses means that on the average the errors involved in such predictions are smaller than those involved in pure guesses by the given per cent. To illustrate this, let us suppose that if pure guesses are made in a certain situation they involve one error of 15 points, one of 14, two of 13, two of 12, and so on. With predictions 40 per cent better than pure guesses the errors would be 40 per cent smaller, so that it might be expected that instead of the error of 15 points there would be one of 9 points, instead of the one of 14 points there would be one of 8.4 points, likewise two of 7.8 points, two of 7.2 points, and so on. Sometimes the size of the errors is expressed in just the opposite way to that so far used in this paragraph, that is to say, instead of saying that predictions are 40 per cent better than pure guesses one may say that they are 60 per cent pure guesses or that the errors involved are 60 per cent as large as those in pure guesses.

Table V shows how the predictions based on coefficients of correlation of given sizes compare with pure guesses. In this table, the first column contains values of the coefficient of correlation at intervals of .01 from 1.00 down to .95 and at intervals of .10 from .90 down to .00, with the coefficient of alienation or fraction of a pure guess correspond-

[2]In actual practice a coefficient of exactly zero will rarely be obtained because with a small number of cases the element of chance agreement or disagreement between prediction and actual facts is fairly large.

TABLE V. COEFFICIENTS OF ALIENATION* CORRESPONDING TO CERTAIN VALUES OF THE COEFFICIENT OF CORRELATION

Coefficient of Correlation	Coefficient of Alienation	Coefficient of Correlation	Coefficient of Alienation
1.00	.0000	.70	.7141
0.99	.1411	.60	.8000
0.98	.1990	.50	.8660
0.97	.2431	.40	.9165
0.96	.2800	.30	.9539
0.95	.3122	.20	.9798
0.90	.4359	.10	.9950
0.80	.6000	.00	1.0000

*The coefficient of alienation is obtained by solving $\sqrt{1-r^2}$, in which r is the symbol for the coefficient of correlation.

ing to each. Beginning at the top of the table it will be seen that when the correlation is perfect and the coefficient 1.00, the coefficient of alienation is zero, or, in other words, prediction can be made with absolute accuracy. When the correlation is .99, the prediction is about .14 of a pure guess, when it is .98 the prediction is .20 of a pure guess, and so on. One can see that the inaccuracy of prediction or the fraction of a pure guess involved therein increases very rapidly at first for comparatively small decreases in the coefficient of correlation. By the time the latter reaches .80, the errors in predictions are .60 as large as those in pure guesses, and when the correlation is .50 the errors are almost .87 as large as those in pure guesses.

If one now recalls the sizes of the coefficients of correlation between freshman marks and other criteria, and then the corresponding coefficients of alienation, he will see at once how unreliable are the best predictions of college marks which can be made upon the basis of these criteria. Most of the obtained simple coefficients of correlation were below .50, very few rising above this. Thus for most of them the best predictions possible are 87 per cent or more pure guesses. A very few of the simple coefficients, and not a great many of the multiple ones, rose above .60. For one of .60 the errors in prediction are .80 as large as those in pure guesses and for one of .65, not given in the table, they are .76 as large. The general statement may therefore be made that, with one or two possible exceptions, the best predictions possible from the criteria used in this study are still subject to errors which are on the whole three-fourths as large as those in pure guesses and that in most cases they are at least four-fifths or five-sixths as large. In other words, they are so large that probably the most that can be said for predictions based upon these or similar criteria is that if it is necessary

or highly advisable to make selection or classification of some sort, these criteria furnish a somewhat better basis for doing so than would mere guesses.

Interpretation of the coefficient of correlation in terms of the probable error of estimate. Another means of describing the accuracy of prediction based on a coefficient of correlation of a given size is to state the probable error of estimate.[3] The probable error of estimate is easily obtained after the coefficient of alienation has been found as all that is necessary is to multiply the latter by the median deviation[4] of the distribution in question. The meaning of the probable error of estimate is, in a general way, the same as that of any other probable error or median deviation, that is, half of the errors involved in making estimates or predictions are less than the probable error and half are greater, about 82 per cent are less than twice the probable error and 18 per cent greater, almost 96 per cent less than three times the probable error and slightly over 4 per cent greater, and so on. For example, a probable error of four points in connection with estimates of freshman marks in algebra would mean that half of the estimates or predictions of the marks made by individual students would be in error by less than four points and half by more, that about 82 per cent of them would be an error by less than eight points and about 18 per cent by more, and so on. The form of statement may be changed to read that the chances are even that the error in the case of any particular individual is not greater than four points, that they are about 4.6 to 1 that it is not greater than twice this amount or eight points, 22 to 1 that it is not greater than three times the probable error or twelve points, and so on.[5]

It will be recalled that the probable errors of estimate given in Chapter V as corresponding to the highest multiple coefficients of correlation obtained were mostly between four and six points, though one or two were slightly smaller than four points and several larger than six, one even being above seven. The central tendency was somewhat above five. In other words, on the average, predictions of college freshman

---

[3]The standard error of estimate might also be used, but the discussion will be confined to the probable error because it is probably more generally understood and used. The standard error is 1.4826 times the probable error.

[4]Since the median deviation equals .6745 times the standard deviation, the usual formula for the probable error of estimate is $.6745\sigma\sqrt{1-r^2}$, in which σ is the abbreviation for the standard deviation.

[5]In case the standard error has been used instead of the probable error the interpretation along the lines given above must, of course, be appropriately changed. The proper per cents and chances for this purpose may be found in Odell, *op. cit.*, p. 14.

TABLE VI. APPROXIMATE DISTRIBUTIONS OF COLLEGE FRESHMAN MARKS AS COMPARED WITH PREDICTIONS THEREOF CORRESPONDING TO VALUES OF THE COEFFICIENT OF CORRELATION OF .40, .50, .60 and .70

r = .40

Criterion Rating	Freshman Mark					
	E	D	C	B	A	T
A		1	3	3	3	10
B	1	2	8	6	3	20
C	3	8	18	8	3	40
D	3	6	8	2	1	20
E	3	3	3	1		10
T	10	20	40	20	10	100

r = .50

Criterion Rating	Freshman Mark					
	E	D	C	B	A	T
A			3	3	4	10
B		4	7	6	3	20
C	3	7	20	7	3	40
D	3	6	7	4		20
E	4	3	3			10
T	10	20	40	20	10	100

r = .60

Criterion Rating	Freshman Mark					
	E	D	C	B	A	T
A			2	3	5	10
B		3	7	7	3	20
C	2	7	22	7	2	40
D	3	7	7	3		20
E	5	3	2			10
T	10	20	40	20	10	100

r = .70

Criterion Rating	Freshman Mark					
	E	D	C	B	A	T
A			1	3	6	10
B		2	7	8	3	20
C	1	7	24	7	1	40
D	3	8	7	2		20
E	6	3	1			10
T	10	20	40	20	10	100

marks in particular subjects based upon the combination of criteria which gave the best prediction in each case would contain errors of more than five points in about half of the cases, of almost eleven points in about 18 per cent, and of about sixteen points in over 4 per cent. The predictions of the general freshman average, which had the smallest probable error of estimate, would be in error by almost four points or more in half of the cases, by about seven and one-half or more in 18 per cent of the cases, and by over eleven points in 4 per cent of the cases. In view of the fact that the total range of passing marks is only 30, it can be seen how serious errors of this size are and how little reliability such predictions possess. On the average about one-half of the students who would really receive freshman marks of 75 would be rated below 70 or as failing, about 18 per cent of those who made 80 would be rated as failing and about 4 per cent of those making 85 would be so rated. Similar per cents would of course be rated too high as well as too low. It is, therefore, evident that a great amount of individual injustice would be done.

Interpretation of the coefficient of correlation in terms of the frequency and amount of displacement. Another method of making

TABLE VII. SUMMARY OF THE AMOUNT OF AGREEMENT AND DISAGREEMENT BETWEEN PREDICTIONS AND ACTUAL FRESHMAN MARKS SHOWN IN TABLE VI

r	Number of Divisions Displaced					
	0	1	2	3	4	T
.70	52	40	8	0	0	56
.60	46	40	14	0	0	68
.50	40	40	20	0	0	80
.40	36	44	16	4	0	88

more concrete the meaning of predictions based upon coefficients of correlation is to compute the numbers of individual cases which would not be correctly predicted if classified in a few groups on the basis of probable academic success. Since what is probably the most common system of marking in college makes use of five letters and since many high schools likewise do the same, the interpretations to be given will be based upon this number of divisions. That is to say, it is assumed that high-school marks and other criteria employed are represented by only five marks and college success likewise by five. It is further assumed that the distribution of marks in each case is such that 10 per cent each of the highest and lowest marks are given, 20 per cent each of the next to the highest and next to the lowest and 40 per cent of average marks. For purposes of convenience these marks are called A, B, C, D, and E, A being the highest and E the lowest. The interpretations given are for coefficients of correlation of .40, .50, .60 and .70, as this range includes practically all of the highest ones obtained in this study. The most likely symmetrical distribution in each case is shown in Table VI. The upper left-hand quarter of the table, for a coefficient of .40, shows that of ten individuals for whom the criterion rating or prediction was A, only three would on the average actually receive A's, three B's, three C's, and one D. For the twenty for whom B's were predicted three would get A's, six B's, eight C's, two D's and one E, and so on. Looking through all four tables it is readily seen that as the correlation increases the number of exact agreements between prediction and fact becomes closer. Thus when $r = .50$, four of the ten whose predicted standing is A actually receive that mark, for $r = .60$, five actually receive it, and for $r = .70$, six.

Although such tables as those above give a rather concrete idea of the situation, yet because of the fact that each contains many entries it is rather hard to get a general or summary idea thereof. This is perhaps

better accomplished by taking one more step and tabulating, as in Table VII above, the numbers of cases out of the hundred in which the agreement is perfect, the number in which there is one step disagreement, two steps disagreement and so on. The bottom line of the table for $r = .40$ corresponds to the upper left-hand fourth of Table VI. It shows, first, that 36 of the hundred cases agree exactly. This number is obtained by adding the three cases which it is predicted will and which do receive A's, the six which are predicted to and do receive B's, the eighteen similar C's, the six D's and the three E's, or in other words by adding the five entries lying on a diagonal line from the upper right-hand to the lower left-hand corner. Furthermore, 44 of the one hundred cases show an error or displacement of one division between predictions and actual marks. This is obtained by taking the sum of the two diagonal rows next to the one from corner to corner, one being on each side of it. Continuing, there are 16 cases which are displaced two divisions or letters and 4 which are displaced three. The total amount of displacement, given in the last column, is 88 and is found by multiplying each of the entries in that row by the amount of displacement under which it falls and adding the products. Thus $44\times1+16\times2+4\times3=88$. It will be seen that even if $r = .70$, which is a much higher correlation than was obtained in the case of practically any of the freshman subjects dealt with in this study, only slightly over half of the cases are correctly predicted and that the total displacement amounts to a shift of 56. The figures for $r = .50$ and .60 more nearly portray the situation existing in most subjects and show that in less than half of the cases would accurate predictions be made, that in about two-fifths of them predictions would be in error by one division or letter, and in one-fifth or less by two.

Summary of this chapter. Since statements of the closeness of relationship between freshman marks and various other items of information do not indicate very concretely the size of the errors involved in predicting the former from the latter when made in terms of coefficients of correlation, several other means have been used. These are the coefficient of alienation, the probable error of estimate, and a statement of the frequency and amount of displacement. It appears that with one or two rather doubtful exceptions the best predictions possible from the available criteria have coefficients of alienation of three-fourths or larger, or, in other words, the errors present are at least three-fourths as large as they would be if pure guesses were made. The corresponding probable error of estimate is not far from five points, which means that

half of the predictions would be in error by more than this amount. Computing the displacement on the assumption of five divisions in both criteria and freshman marks the best predictions would result in the displacement of about half of the individuals by one or two divisions. Thus on the whole it is apparent that the errors present in the best predictions are both numerous and large and that if the predictions are used comparatively little confidence should be placed in them.

CHAPTER VII

IS THE CHANGE FROM HIGH SCHOOL TO COLLEGE GREATER THAN THAT FROM ELEMENTARY TO HIGH SCHOOL?

A further statement of the question. A problem which has not infrequently been discussed, especially since the public high school has become such an important part of our educational system, is that of the breaks between the elementary and the high school and between the high school and the college. It has frequently been claimed and the claim fairly well supported that too great a break occurs at one or both places and that those in school are subjected to too abrupt a transition. It occurred to the writer, therefore, that it might be a matter of interest to present very briefly a few comparative data which would tend to show which, if either, of the two changes was the greater.

The data to be used in the comparison. One means of determining how great the change is from one type of school to another is by means of the correlation between marks before and after the transition. Thus the correlations obtained between high-school and college freshman marks in this study are in a sense measures of the similarity of work done by the same individuals in high school and college and of the conditions under which this work is done. Likewise correlations between elementary-school and high-school marks may be considered as similar measures in that case. Although such correlations are in neither case perfect or entirely satisfactory measures yet they are in both cases subject to practically the same limitations and therefore may be fairly compared with each other.

Instead of making an exhaustive, or even fairly wide, study and compilation of the various correlations obtained between elementary and high-school marks, the writer has selected a single study[1] along this line and will compare the results given in it with those obtained in his own study. This study was selected because it appears to be one of the most carefully conducted investigations dealing with this problem which has been carried out and also includes data for a number of groups of pupils.

---

[1]Ross, C. C. "The relation between grade school record and high school achievement." Teachers College Contributions to Education, No. 166. New York: Teachers College, Columbia University, 1925. 70 p.

TABLE VIII. COEFFICIENTS OF CORRELATION BETWEEN MARKS IN ELEMENTARY AND HIGH-SCHOOL FRESHMAN SUBJECTS*

Elementary School Subject	High-School Freshman Subject		
	Average	English	Mathematics
Arithmetic	.52	.44	.38
English	.59	.50	.34
Fine Arts	.35	.48	.08
Geography	.50	.50	.28
History	.40	.40	.24
Reading	.35	.46	.18
Spelling	.36	.45	.18
Special Sub.	.26	.43	.03

*Taken from Ross, op. cit., p. 15.

TABLE IX. COEFFICIENTS OF CORRELATION BETWEEN ELEMENTARY-SCHOOL AVERAGE AND HIGH-SCHOOL FRESHMAN MARKS*

	High-School Freshman Subject			
	Average	English	Latin	Mathematics
New Rochelle, 1916.	.68	.60	.58	.42
New Rochelle, 1917.	.67	.67	.73	.51
New Rochelle, 1918.	.56	.67	.57	.43
New Rochelle, 1919.	.65	.60	.64	.51
Des Moines	.69	.61	.61	.51

*Taken from Ross, op. cit., p. 35.

Tables VIII and IX present some of the correlations which Ross obtained. In the first will be found those between marks in eight elementary-school subjects or groups of subjects and the high-school freshman average, English mark and mathematics mark. The second gives those between elementary-school average and high-school freshman average, English, Latin, and mathematics marks. It will be seen that the coefficients given in Table VIII run from very near zero up to slightly above .50, those of the elementary-school subject marks with mathematics being decidedly lower than those with freshman average and English marks, which on the whole do not differ greatly. The central tendency of the latter, that is, of the correlations with both freshman average and English mark is somewhat above .40, whereas that for mathematics mark is not far from .20. In the case of the elementary-school averages given in Table IX the correlations are distinctly higher and also more uniform. For the freshman average and English and Latin marks they are practically all in the sixties, but in mathematics again somewhat lower.

Comparison of the results of Ross and of the writer. When these are compared with the coefficients given in Table III and discussed in Chapter IV at least two facts are evident. Those in Table VIII on the whole run very much the same as do the corresponding ones in Table III. The correlations of the several elementary subject marks with the freshman average tend to be about the same as those of the particular college freshman subject marks with the high-school average and likewise those of the elementary ones with English and mathematics marks fall in about the same range as those of the particular high-school and college freshman subject marks. The fact that several of those for mathematics in Table VIII are lower than any corresponding ones in Table III is undoubtedly due to the fact that the correlations between subject marks found in the writer's study and given in Table III are only those which it seemed likely would be fairly high, whereas no such selection was employed in finding those in Table VIII. The second noticeable fact is that the coefficients in Table IX tend to be decidedly higher than the corresponding ones in Table III. The correlation between high-school average and college freshman average was only .55, whereas the lowest one between elementary-school and high-school freshman average is .56 and their average .65, .10 greater than that between high-school and college. Those between the high-school average and the college freshman marks in particular subjects were only in two or three cases as high as .60 and more often below .50 than above, whereas eight out of the 15 given in Table IX are .60 or above and only two below .50. On the whole the difference in this case is probably best represented by at least .15 rather than .10.

Conclusion. The evidence just presented supports the conclusion that the break between elementary school and high school is not as marked as that between high school and college. Inasmuch as it is probably true that there are greater differences in the subject matter of elementary-school and high-school subjects than in that of high-school and college subjects, the higher correlations in the former case imply that other factors therein must both balance this greater difference in subject matter, and also indicate greater similarity otherwise. The chief factor is probably that elementary and high schools are usually united and under the same general control within particular school systems, and that both their aims and methods of instruction tend to be more similar than is true in the case of high schools and colleges. In most high schools pupils are subjected to fairly close supervision and their work consists largely in performing very definitely specified assignments under more

or less guidance from teachers, whereas in college very little control is usually exercised over their study habits and assignments are often much more indefinite. In other words, it seems that in high school as well as in elementary school the prevailing practice is to treat the pupils as relatively irresponsible children, whereas in college they are expected to assume full individual responsibility.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

The problem. The very marked recent increase in college enrollment has had as one of its results a considerable growth of interest in the problem of selecting from applicants for admission to college those who will be able to do satisfactory work therein, and therefore in methods of predicting scholastic success in college. Most of the investigations along this line have dealt with making such predictions on the basis of high-school marks, college entrance examinations over high-school subjects, and intelligence test scores.

Results secured by other investigators. The results obtained by a rather large number of investigators may be summarized by saying that there is probably little difference in the accuracy of predictions based upon these three different criteria and that most of the coefficients of correlation between any one of them and college marks may be expected to range from about .40 to .50 or perhaps higher; that combinations of these factors will often yield correlations of .60 or above; and that if measures of study and other relevant habits are included this figure can be raised appreciably.

The writer's results. A study conducted by the writer which included almost two thousand college freshmen in over a hundred different institutions did not deal with ordinary entrance examinations or with study habits. The correlations found between high-school and freshman marks were about the same as those obtained by other investigators and those between test scores and freshman marks somewhat lower, undoubtedly due to the fact that the test used was not the most reliable one available for this purpose. When multiple correlation was employed comparatively small increases were produced in the coefficients, only a few of them rising above .60 when the best combinations of school marks and test scores were made. This was largely due to the fact that many of the high-school marks used in simple correlation were averages of several subjects. In this connection it should be noted that the correlations obtained in the writer's study were undoubtedly lowered somewhat because of the throwing together of school marks from hundreds of different high schools and colleges which must have introduced errors that lowered the coefficients.

Accuracy of predictions based on obtained data. Coefficients of correlation of the sizes usually found, that is from .40 up to .50 or even .60, indicate that the corresponding predictions of freshman marks are not a great deal better than pure guesses. Predictions corresponding to the highest of these coefficients are still subject to errors at least three-fourths as large as those in pure guesses or, in terms of points on the percentile marking scale, at least half of the errors in the estimates are larger than five points. It should, however, be pointed out that high scores upon intelligence tests and also, but probably to a lesser degree, high high-school or entrance examination marks are more reliable than low ones and more confidence may be placed in them as indicating possibilities of student performance. Comparatively few individuals earn scores or marks much above what they really deserve, but many fairly able individuals, through indifference, carelessness, temporary distraction or other causes, make scores or marks too low to be indicative of the abilities which they possess. One is reasonably safe, therefore, in assuming that an individual who is rated high by all, or even by any one, of the three criteria—intelligence test, high-school mark, and college entrance examinations—is able to do reasonably satisfactory or better work in college, whereas the fact that an individual's score is low in one of them or perhaps even in all is not nearly so sure an indication that he cannot, if he will, do satisfactory work, although the chances are strong that if his scores are low in all three, his scholastic work in college will also be decidedly poor.

Comparative break between elementary and high school and between high school and college. A comparison of the results of this study with those obtained by Ross in predicting high-school from elementary-school marks shows that the correlations in the latter case are distinctly higher, the differences averaging .10 or .15. In other words, it appears that the similarity of the work done or, more likely, of the conditions prevailing, in elementary and high school is greater than it is in the case of high school and college.

Conclusion as to the use of the available criteria for predicting scholastic success in college. Since both for the purpose of determining the admission of applicants to college and of personnel work with students who have been admitted, it is desirable to predict the quality of their scholastic work, it is better to make use of such bases of prediction as we have than to rely upon pure guesses. Therefore the use of high-school marks, entrance examination results, or intelligence test scores for this purpose is better than that of no criterion at all. It

should be recognized, however, that the errors present are decidedly large and that the predictions of the work to be done by individuals cannot be relied upon as possessing a high degree of accuracy. The writer wishes to emphasize as strongly as possible the need for further experimentation and investigation, especially along the line of determining which criteria form the best combinations for predictions based upon multiple correlation and regression equations. He believes that a combination of the score upon one of our best intelligence tests for this purpose, such as that of Thorndike, of the marks in certain high-school subjects or groups of subjects and on the best types of entrance examinations over these subjects, of the ratings of study habits and perhaps other factors, can be combined so that practically any institution can obtain correlations of .75 or .80 between the marks of its students and the best available combination.